A novel *Bacillus natto* plasmid pLS32 capable of replication in *Bacillus subtilis*

Teruo Tanaka*, Mitsuo Ogura

Department of Marine Science, School of Marine Science and Technology, University of Tokai, Orido 3-20-1, Shimizu, Shizuoka 424, Japan

Received 2 December 1997; revised version received 19 December 1997

Abstract Plasmid pLS32 is a relatively large (approximately 70 kbp), cryptic, low copy-number plasmid present in *Bacillus natto*. We isolated and analyzed the replication region of the plasmid in *B. subtilis*, and the following results were obtained: the replication region contained an open reading frame encoding a 287-amino acid protein (RepN), whose amino acid sequence was partially homologous with those of the Rep proteins encoded on plasmids pAD1 and pLJ1 isolated from *Enterococcus faecalis* and *Lactobacillus helveticus*, respectively; the replication origin (*oriN*) was located in the *repN*-coding region; the copy number of a pLS32 derivative, pBET131, was 2 to 3 per chromosome; replication of pBET131 required *polI*. These features show that pLS32 is a novel plasmid capable of replication in *B. subtilis*.

Key words: *Bacillus natto* plasmid; Copy number; *polI* independence; *repN*

1. Introduction

Plasmids capable of replication in *Bacillus subtilis* are mostly derived from heterologous Gram-positive bacteria, and they are divided into two classes in terms of their mode of replication. One class includes plasmids using the rolling circle mode of replication, and these are present at relatively high copy numbers per cell. The *Staphylococcus aureus* plasmids pUB110, pC194, and pE194 are representatives of this group [1]. Plasmids belonging to the other class replicate in the host through a theta type intermediate, and these are present at low copy numbers. These plasmids are derived from various Gram-positive species such as *Enterococcus faecalis* (pAMβ1), *Streptococcus agalactiae* (pIP501), *Streptococcus pyogenes* (pSM19035), *Clostridium perfringens* (pIP404), and *Bacillus stearothermophilus* (pTB19) [2].

In addition to those described above a series of homologous plasmids exist, which are derived from various strains of *B. subtilis* and *B. natto*, a species closely related to *B. subtilis* [3–7]. They are also divided into two classes based on their sizes and mode of replication. The small plasmids are present in the host at high copy numbers, and several of them were shown to replicate by the rolling circle mechanism [8]. The large plasmids are present at low copy numbers [3,5], and one of them, pLS20, was recently suggested to replicate via a theta type intermediate [9]. Plasmid pLS32, for which no apparent phenotype is known, has a molecular size of about 70 kbp and belongs to the latter class of *B. natto* plasmids. It has recently been demonstrated that the replication region of pLS32 can support replication of a DNA fragment as large as 310 kbp without gross DNA rearrangement [10], and even of the entire chromosome of *B. subtilis* via bidirectional replication [11]. These unique features suggest that the plasmid is useful as a vehicle for cloning large DNA sequences in *B. subtilis*, and prompted us to characterize the plasmid further.

We describe in this report several characteristics of the replication region of pLS32.

2. Materials and methods

2.1. Strains and plasmids

B. subtilis 1A224 (*hisH2 pheA1 trpC2*), and its *polI* mutant 1A226 were obtained from the Bacillus Genetic Stock Center (Columbus, OH). *B. subtilis* MI112 *arg-15 leuB8 thr*-5 *recA4 hsmM hsrR,* TT718 *trpC2 leuC7 degS*::*cat* and *B. natto* IAM1163 were described previously [5,12,13]. A *B. natto* plasmid, pLS32, was isolated as described previously [5].

2.2. Construction of pLS32 derivatives

Digestion with *Bam*HI and ligation of both pLS32 and a vector pUH101 [14], a pBR322 derivative carrying the chloramphenicol resistance gene (*cat*), followed by transformation into *B. subtilis* MI112 gave rise to plasmid pB1. The *Bam*HI insert (7.4 kbp) was excised, partially digested with *Hin*dIII, and cloned into a pBR322 derivative carrying the *cat* and tetracycline resistance (*tet*) genes (unpublished), resulting in plasmid pBET131. The *Bam*HI-*Hin*dIII insert (7.2 kbp) in pBET131 (Fig. 1) was further deleted to construct plasmids pSEQ243, 131, 231, 2426 by cloning appropriate restriction fragments into an *E. coli* plasmid pUKM504 carrying the neomycin resistance (*neo*) gene [15]. Plasmid pSEQ24310 was constructed by cleavage of pSEQ243 with *Acc*I, followed by blunting with T4 DNA polymerase and ligation. Insertion of the *cat*-containing fragment derived from pBEST4C [16] between the two *Hpa*I sites of pSEQ243 gave rise to pSEQ24308. Plasmid pHDDS3 was constructed by insertion of the *Alw*NI-*Hin*dIII fragment downstream of an IPTG-inducible promoter P_{spac} in pDH88 [17]. Plasmids pHDCS2, pHDSD13, and pHDNS2 were constructed by insertion into *Hin*dIII- and *Bgl*II-treated pDH88 of PCR-amplified DNA fragments which had been digested with *Hin*dIII and *Bam*HI. The following PCR primers were designed so that the 5′ and 3′ ends of the amplified products carry *Hin*dIII and *Bam*HI sites, respectively, using the nucleotide sequence shown in Fig. 2. Primers 3 (5′-GTAAGCTTCTGAGCAAACCGACCAAAGT-3′ containing nucleotides (nt) 6137 through 6156) and 2 (5′-AAGGATCCAGCTTCCTCCTAAAATCTTT-3′ containing nt 7228 through 7209), primers 1 (5′-GTAAGCTTAAAGGAGCAACAAAAATGAG-3′ containing nt 6241 through 6260) and 4 (5′-AAGGATCCTTAATTATCTAACCAATTATAAA-3′ containing nt 7119 through 7097), primers 5 (5′-GTAAGCTTGGAGGTTGTTTTATGAGTAAATATTTCACAGC-3′ containing 6256 through 6275; the underlined sequences indicate the SD sequence and initiation codon, respectively) and 4, and primers 6 (5′-GTAAGCTTATGAGTAAATATTTCACAGC-3′ containing nt 6256 through 6275; the initiation codon is underlined) and 4 were used for construction of pHDDS3, pHDCS2, pHDSD13, and pHDNS2, respectively.

2.3. DNA sequence

The nucleotide sequence data reported in this paper will appear in the GSDB, DDBJ, EMBL, and NCBI nucleotide sequence databases under the accession number D49467.

*Corresponding author. Fax: (81) (543) 34 9834.
E-mail: teruot@scc.u-tokai.ac.jp

PII S0014-5793(98)00015-5

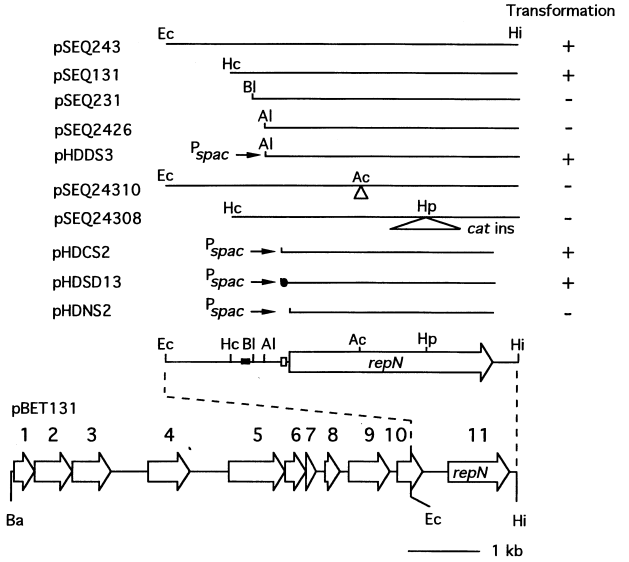

Fig. 1. Localization of the region containing the replication function. The open arrows indicate the direction and sizes of the potential open reading frames. A 1.5-kbp region containing the *repN*-coding sequence is expanded between the two dotted lines. *cat* ins shows that the *cat* gene has been inserted at the restriction site shown. The open triangle in pSEQ24310 depicts a frame shift mutation introduced at the *Acc*I site. Symbols + and − show the transforming ability of each plasmid. The dark and open boxes show the putative promoter (the −35 and −10 sequences) and the SD sequence of *repN*, respectively. The filled circle in pHDSD13 shows the synthetic SD sequence. Construction of the plasmids is described in Section 2. The following abbreviations were used: Ac, *Acc*I; Al, *Alw*NI; Ba, *Bam*HI; Bl, *Bgl*I; Ec, *Eco*RI; Hc, *Hin*cII; Hi, *Hin*dIII; Hp, *Hpa*I.

3. Results

3.1. Isolation of a DNA fragment carrying the replication function of pLS32

The replication and its flanking regions of pLS32 were isolated in a 7.2-kbp fragment of pBET131 as described in Section 2, and the deduced open reading frames contained in the fragment are shown in Fig. 1. The rightmost 1.5-kbp *Eco*RI-*Hin*dIII region contained an autonomous replication ability as shown by subcloning experiments (Fig. 1). It carried a potential ORF, designated *repN* that could encode a protein of 287 amino acids. The *repN* ORF is preceded by a putative SD sequence located between nt 6241 and 6247 (Fig. 2). A homology search in the PIR and SWISS-PROT databases with the FASTA program revealed that the RepN protein had 29.4% identity over a 218-amino acid overlap to the RepA protein of pAD1 isolated from *E. faecalis* [18], and 25.9% identity over a 224-amino acid overlap to the 41-kDa protein of a *Lactobacillus helveticus* plasmid pLJ1 [19]. The 41-kDa product is likely to be a Rep protein, since it is the only protein encoded on pLJ1. In both cases the homology was high at the N-terminal regions of the proteins compared.

To further delineate the DNA region necessary for autonomous replication, derivatives of pSEQ243 were constructed and tested for transforming activities (Fig. 1). Deletions extending to both the *Bgl*I (pSEQ231) and *Alw*NI (pSEQ2426) sites but not to the *Hin*cII site (pSEQ131) abolished the transforming activity. When the region downstream of *Alw*NI was placed under the control of the P_{spac} promoter (pHDDS3), the transforming activity was restored, suggesting that the failure of pSEQ231 and pSEQ2426 to transform the host strain is

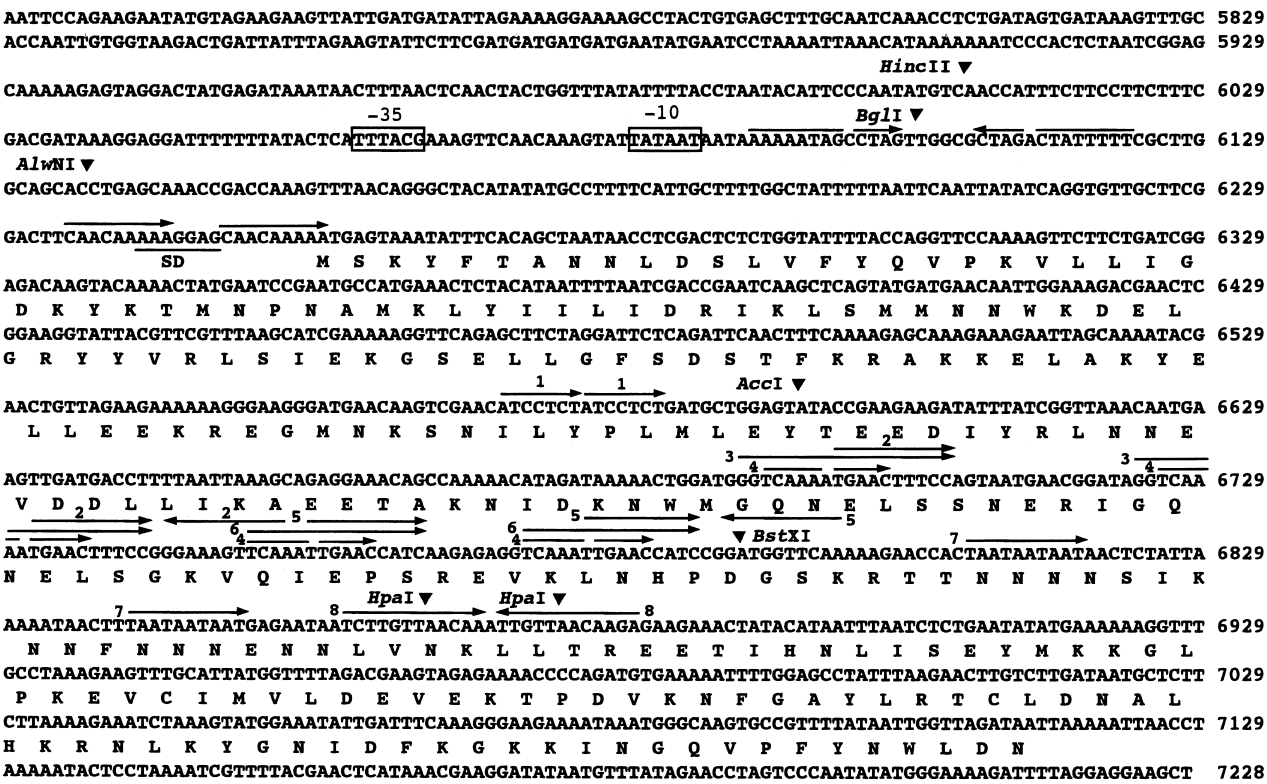

Fig. 2. Nucleotide sequence of the 1.5-kbp *Eco*RI-*Hin*dIII fragment containing the replication region. The nucleotides are numbered from the *Bam*HI site (see Fig. 1). The boxed and underlined sequences represent the potential promoter and SD sequences, respectively. Arrows indicate direct and inverted repeat sequences, and those with the same number show the sequences having the same nucleotide sequences.

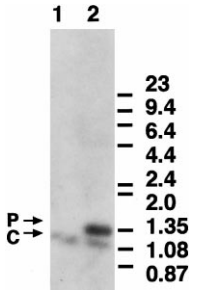

Fig. 3. Estimation of the copy number of pBET131 by Southern hybridization. A digoxigenin-labeled PCR fragment containing the *cat* gene was hybridized with DNAs from TT718 (lane 1) and from TT718 carrying pBET131 (lane 2). The arrows marked by P and C show that the DNA fragments were derived from the chromosomal DNA and plasmid DNA, respectively. The lines and numbers on the right indicate the positions and sizes (kbp) of the DNA size markers, respectively. The intensity of the hybridized bands on X-ray film was quantitated by densitometric analysis using Shimadzu CS-930PC.

due to the lack of a promoter present in the 97-bp region between the *Hin*cII and *Bgl*I sites. A computer search for a possible promoter in this region revealed a sigma A type promoter [20] consisting of two hexamers 5′-TTTACG-3′ and 5′-TATAAT-3′ expected for the −35 and −10 regions, respectively, with an interval of 17 nucleotides (Fig. 2). A frame shift mutation introduced at the *Acc*I site in the *repN* sequence (pSEQ24310), and disruption of the ORF by insertion of the *cat* gene in between the two *Hpa*I sites (pSEQ24308) inactivated the transforming activity. Deletion of both the 97 nucleotides from the *Alw*NI site to the nucleotide just before the putative SD sequence and the region downstream of the stop codon of the *repN* gene (pHDCS2) did not affect the transformability of the plasmid.

We next tested if there was a particular sequence between the putative SD sequence and the initiation codon of *repN* that is involved in replication. We changed the SD sequence from 5′-AAAGGA-3′ to 5′-GGAGG-3′ and its immediately downstream sequence from 5′-CAACAAAA-3′ to 5′-TTGTTTT-3′ (see Fig. 2, and Section 2). Plasmid pHDSD13 thus constructed was found to have transforming activity, indicating that the nucleotide sequence before the initiation codon does not contain the information necessary for replication. We further constructed pHDNS2 which contained the entire *repN*-coding sequence but lacked the SD sequence (see Section 2), and found that the plasmid was unable to transform the host strain. These results, together with those with disruption of *repN* (pSEQ24310 and pSEQ24308), show that the deduced RepN protein is necessary for replication, and that the origin of replication (*oriN*) is within the *repN*-coding region.

3.2. Copy number

We studied by Southern hybridization the copy number of pBET131 by comparing the amounts of the *cat* gene carried on pBET131 and the chromosome of *B. subtilis* TT718 in which one copy of the *cat* gene had been inserted at the *degS* gene [13]. In the TT718 chromosome, one *Xba*I site is present 100 bp before the initiation codon of *degS* and another downstream of the inserted *cat* gene [13], while in pBET131, the *cat* gene is flanked by *Hin*dIII and *Bam*HI sites separated by 1.3 kbp. Consequently, digestion of DNAs from TT718 and TT718 carrying pBET131 with *Xba*I, *Hin*dIII, and *Bam*HI gives rise to a 1.2-kbp and both 1.2- and 1.3-kbp fragments containing the *cat* gene, respectively. A typical hybridization experiment showed banding patterns as depicted in Fig. 3. A densitometric analysis of the bands in the negative showed that the intensity of the *cat* region derived from pBET131 was 2.1-fold as compared to that of the *cat* region from the chromosome (Fig. 3, lane 2). In the culture of TT718(pBET131) from which the DNA was prepared, the fraction of the cells resistant to both Tc and Cm (the resistance markers on pBET131) was 95%. Therefore, by correcting for the cells that had lost the plasmid, the copy number of pBET131 was calculated to be 2.2/chromosome. From several independent experiments the copy number was estimated to be 2–3/chromosome.

3.3. Replication of oriN-containing pBET131 in polI strain

Some other low copy number plasmids showing theta type replication, such as pAMβ1, pIP501 and pSM19035, require *polI* for replication [21]. We found that pBET131 transformed both the 1A224 (*polI*$^+$) and 1A226 (*polI*) cells (data not shown), indicating that PolI is not necessary for the replication of *oriN*-driven plasmids.

4. Discussion

The results described in this study show that the replication origin (*oriN*) of *B. natto* plasmid pLS32 is located within an open reading frame encoding a putative replication protein RepN, that the copy number of pBET131 carrying the pLS32 replication region is 2–3/chromosome and that replication does not require PolI.

Single-stranded DNA was not detectable in the cell carrying a pLS32 derivative (data not shown), and no characteristic features which are observed around the single strand origin (SSO) region [1,8,22,23] were found in the *repN*-coding region. It has been demonstrated that the replication of the *B. subtilis* chromosome driven by the replication machinery of pLS32 proceeds bidirectionally from the *oriN*-containing site [11]. A DNA fragment as large as 300 kbp was shown to be maintained by pBET131 [10], in contrast to the observations that structural alterations are often observed when RCM plasmids are used as a cloning vehicle [1,22]. These properties of the pLS32 replicon strongly suggest that it uses the theta mechanism of replication.

It was shown by deletion experiments (Fig. 1) that minimum requirements for autonomous replication were within the *repN*-coding sequence. One remarkable feature of the *repN*-coding region is that there are 7 direct (numbered 1 through 7 in Fig. 2) and 3 inverted (numbered 2, 5 and 8) repeat sequences. The sequences 2 and 5 which constitute both direct and inverted repeats are present in direct repeats 3 and 6, respectively. The long direct repeats 3 and 6 are related to each other as shown by the presence of the direct repeat 4 in their sequences. It is possible that some of the repeats are involved in *oriN* function, since it has been demonstrated for *E. coli* plasmids R6K [24], P1 [25], F [26] and pSC101 [27] that a plasmid-encoded protein recognizes direct repeats and regulates plasmid replication, and it has been suggested that direct or inverted repeats are involved in the *ori* function for a *B. natto* plasmid pLS20 [9].

Low copy number plasmids that use the theta mechanism

of replication can be classified into five groups, classes A to E [9,21]. The plasmids belonging to classes A, C, and E encode Rep proteins, and among these the class C and E plasmids require PolI for replication. The class A plasmids have a characteristic replication origin called *oriA* which contains repeated sequences (iterons), DnaA boxes and an AT-rich region with repeats. Replication of pLS32 requires RepN but not PolI as shown in this study. Furthermore, the *repN* sequence contains neither the DnaA box nor a noticeable AT-rich region containing repeats (Fig. 2), and in contrast to the replication regions of the class A plasmids, *oriN* is within the coding region of the *repN* gene. These features indicate that pLS32 can be classified as a new class belonging to Gram-positive bacteria. We note that the *B. natto* plasmid pLS20 is different from pLS32 in that the former does not encode a Rep protein [9].

The replication origin region of pLS32 will be useful to clone large DNA molecules as exemplified by the successful cloning of a 300-kbp DNA fragment in pBET131 [10], and could lead to construct a cell carrying an 'artificial chromosome'.

Acknowledgements: We thank M. Itaya for plasmids, and D. Ziegler for bacterial strains. We also thank S. Suzuki for technical assistance. This work was supported in part by a Grant-in-Aid for Creative Basic Research on 'Human Genome Analysis' from the Ministry of Education, Science and Culture of Japan, and a fund from Takano Life Science Research Foundation.

References

[1] Gruss, A. and Ehrlich, S.D. (1989) Microbiol. Rev. 53, 231–241.
[2] Janniere, L., Gruss, A. and Ehrlich, S.D. (1993) in: *Bacillus subtilis* and Other Gram Positive Bacteria: Biochemistry, Physiology, and Molecular Genetics (Sonenshein, A.L., Hoch, J.A. and Losick, R., Eds.) pp. 625–644, American Society for Microbiology, Washington, D.C.
[3] le Hegarat, J.-C. and Anagnostopoulos, C. (1977) Mol. Gen. Genet. 157, 167–174.
[4] Lovett, P.S. and Bramucci, M.G. (1975) J. Bacteriol. 124, 484–490.
[5] Tanaka, T. and Koshikawa, T. (1977) J. Bacteriol. 131, 699–701.
[6] Tanaka, T., Kuroda, T. and Sakaguchi, K. (1977) J. Bacteriol. 129, 1487–1494.
[7] Uozumi, T., Ozaki, A., Beppu, T. and Arima, K. (1980) J. Bacteriol. 142, 315–318.
[8] Meijer, W.J.J., Venema, G. and Bron, S. (1995) Nucleic Acids Res. 23, 612–619.
[9] Meijer, W.J.J., de Boer, A.J., van Tongeren, S., Venema, G. and Bron, S. (1995) Nucleic Acids Res. 23, 3214–3223.
[10] Itaya, M. and Tanaka, T. (1997) Proc. Natl. Acad. Sci. USA 94, 5378–5382.
[11] Hassan, K.M., Moriya, S., Ogura, M., Tanaka, T., Kawamura, F. and Ogasawara, N. (1997) J. Bacteriol. 179, 2494–2502.
[12] Tanaka, T. (1982) in: Microbiology–1982 (Schlessinger, D., Ed.) pp. 15–18, American Society for Microbiology, Washington, D.C.
[13] Mukai, K., Kawata, M. and Tanaka, T. (1990) J. Biol. Chem. 265, 2000–2006.
[14] Tanaka, T. and Kawata, M. (1988) J. Bacteriol. 170, 3593–3600.
[15] Ogura, M. and Tanaka, T. (1996) J. Bacteriol. 178, 216–222.
[16] Itaya, M., Yamaguchi, I., Kobayashi, K., Endo, T. and Tanaka, T. (1990) J. Biochem. 107, 799–801.
[17] Henner, D.J. (1990) Methods Enzymol. 185, 223–228.
[18] Weaver, K., Clewell, D.B. and An, F. (1993) J. Bacteriol. 175, 1900–1906.
[19] Takiguchi, R., Hashiba, H., Aoyama, K. and Ishii, S. (1989) Appl. Environ. Microbiol. 55, 1653–1655.
[20] Moran, C.P. (1993) in: *Bacillus subtilis* and Other Gram-positive Bacteria: Biochemistry, Physiology, and Molecular Genetics (Sonenshein, A.L., Hoch, J.A. and Losick, R., Eds.) pp. 653–667, American Society for Microbiology, Washington, D.C.
[21] Bruand, C., Le Chatelier, E., Ehrlich, S.D. and Janniere, L. (1993) Proc. Natl. Acad. Sci. USA 90, 11668–11672.
[22] Bron, S., Meijer, W., Holsappel, S. and Haima, P. (1991) Microbiology 142, 875–883.
[23] Seery, L. and Devine, K.M. (1993) J. Bacteriol. 175, 1988–1994.
[24] McEachern, M.J., Filutowicz, M. and Helinski, D.R. (1985) Proc. Natl. Acad. Sci. USA 82, 1480–1484.
[25] Chattoraj, D.K., Snyder, K.M. and Abeles, A.L. (1985) Proc. Natl. Acad. Sci. USA 82, 2588–2592.
[26] Trawick, J.D. and Cline, B.C. (1985) Plasmid 13, 59–69.
[27] Vocke, C. and Bastia, D. (1985) Proc. Natl. Acad. Sci. USA 82, 2252–2256.